Available online at www.sciencedirect.com

ScienceDirect

J. Math. Anal. Appl. 344 (2008) 687–698

Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS

www.elsevier.com/locate/jmaa

Monotonicity and best approximation in Orlicz–Sobolev spaces with the Luxemburg norm

Shutao Chen [a,1], Xin He [a,1], Henryk Hudzik [b,*], Anna Kamińska [c]

[a] *School of Mathematics and Computer Science, Harbin Normal University, Harbin 150080, PR China*
[b] *Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Umultowska 87, 61-614 Poznań, Poland*
[c] *Department of Mathematical Sciences, The University of Memphis, Memphis, TN 38152, United States*

Received 9 May 2007

Available online 23 February 2008

Submitted by M. Milman

Abstract

Criteria for strict monotonicity, upper (lower) locally uniform monotonicity and uniform monotonicity of Orlicz–Sobolev spaces with the Luxemburg norm are given. Some applications to best approximation are presented.

Keywords: Orlicz–Sobolev spaces; Strict monotonicity; Upper (lower) locally uniform monotonicity; Uniform monotonicity; Best approximation

The monotonicity properties of lattices have been introduced and studied in the context of their geometric structure [5]. In 1985, M.N. Akcoglu and L. Sucheston [2] showed how the strict and uniform monotonicity were related to ergodic theory. In 1992, W. Kurc [28] observed that the role of monotonicity properties in Banach lattices is similar to the role of rotundity properties in Banach spaces. The relations among rotundity and monotonicity properties in Banach lattices were further studied in [20]. It was noted in [24] that monotonicity properties have close relationships to complex rotundities and their applications. Monotonicity properties have been extensively studied by several authors in specific lattices as Lorentz, Orlicz or Musielak–Orlicz spaces [9,15,19,20,22–24,28,31,32]. For instance in [22], the authors introduced the concept of locally uniform monotonicity and investigated it in Musielak–Orlicz spaces. In [9,31,32] monotonicity and monotone coefficients were discussed in function and sequence Orlicz spaces. The criteria for strict and uniform monotonicity in Lorentz spaces were found in [19].

In this paper we study the monotonicity properties and their applications to approximation theory in Orlicz–Sobolev spaces. Sobolev spaces play very important role in the theory of nonlinear partial differential equations [1]. Their generalizations, Orlicz–Sobolev spaces, have been also used for that purpose (cf. [4,11–13]). Sobolev spaces have been generalized in many different ways, among others to Orlicz–Sobolev or Musielak–Orlicz–Sobolev spaces (cf. [10,14,16–18]).

* Corresponding author.
E-mail addresses: chensht01@yahoo.com.cn (S. Chen), hexin8323@yahoo.com.cn (X. He), hudzik@amu.edu.pl (H. Hudzik), kaminska@memphis.edu (A. Kamińska).

1 Supported by Chinese Science Foundation grant (10471032).

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2008.02.015

It is well known that these properties are applied in approximation theory and in particular they are very useful in estimations of the errors of the approximation [4,21,22,25,28,29]. In [25] the authors present some results on the existence of best approximant in subsets A of a Musielak–Orlicz space L_Φ which are lattice closed, that is, $\sup_n f_n$ and $\inf_n f_n$ are in A whenever $f_n \in A$ for every $n \in \mathbb{N}$, for any $x \in L_\Phi$. For Orlicz spaces the same has been done in [29]. The problem of uniqueness of the best approximant (in usual sense) in Sobolev spaces has been considered in [27].

In this paper we first study some monotonicity properties of Orlicz–Sobolev spaces and then we apply obtained results to the dominated best approximation problems such as the existence, the uniqueness, stability and continuity of the dominated best approximation operator. Let us fix $m \in \mathbb{N}$ and consider the operator

$$Px(t) = \sum_{|\alpha| \leqslant m} a_\alpha \frac{\partial^\alpha x}{\partial t_1^{\alpha_1} \dots \partial t_n^{\alpha_n}}$$

for any $t = (t_1, \dots, t_n) \in \Omega$, $a_\alpha \in \mathbb{R}$, where $\alpha = (\alpha_1, \dots, \alpha_n) \in (\mathbb{N} \cup \{0\})^n$, $|\alpha| = \alpha_1 + \dots + \alpha_n$, and the derivatives of $x \in L_A$ are understood in the distribution sense, that is, $\frac{\partial^\alpha x}{\partial t_1^{\alpha_1} \dots \partial t_n^{\alpha_n}}$ is such a Lebesgue measurable function on Ω, which for any infinitely differentiable in the usual sense measurable function $\varphi : \Omega \to \mathbb{R}$ with compact support in Ω satisfies the equality

$$\int_\Omega \frac{\partial^\alpha x(t)}{\partial t_1^{\alpha_1} \dots \partial t_n^{\alpha_n}} \varphi(t)\, \mathrm{d}t = (-1)^{|\alpha|} \int_\Omega x(t) \frac{\partial^\alpha \varphi(t)}{\partial t_1^{\alpha_1} \dots \partial t_n^{\alpha_n}}\, \mathrm{d}t,$$

where $\frac{\partial^\alpha \varphi(t)}{\partial t_1^{\alpha_1} \dots \partial t_n^{\alpha_n}}$ is the usual mixed derivative of φ.

The dominated best approximation in Orlicz–Sobolev spaces $(W_{m,A}, \|\cdot\|_{m,A})$ can be naturally applied when we are interested in solving the partial differential equation

$$Px(t) = y(t), \quad t \in \Omega, \tag{$*$}$$

where $y \in L_A$ is given and we are looking for all solutions x or generalized solutions $x_0 \in W_{m,A}$, that is $\|Px_0 - y\|_{m,A} = \inf_{x \in W_{m,A}} \|Px - y\|_{m,A}$, of the differential equation $(*)$ satisfying some boundary conditions and/or the dominated condition $x \leqslant z$, where z is some fixed control function from the Orlicz–Sobolev space $W_{m,A}$ and "$\leqslant$" is the partial order defined below.

The lattice approximation in Orlicz–Sobolev spaces investigated here is also related to the constrained interpolation presented for instance in [3, Chapter 10, p. 283]. Lattice approximating of f by elements of K where $K \leqslant f$ in the sense of the partial order considered in Orlicz–Sobolev spaces below, can be also geometrically interpreted as approximation of f by elements from K with smaller oscillation.

In the first section we introduce basic notions, we agree on terminology and provide some results which we will use further in the paper. In the second section we present criteria for uniform monotonicity, upper (lower) locally uniform monotonicity and strict monotonicity of Orlicz–Sobolev spaces. The third section is devoted to applications to lattice best approximation problems in those spaces. We finish this section with a specific example of the convex set $K_m \subset W_{m,A}$ for which we apply the approximation theorems presented in this section.

1. Preliminaries

Let X be a Banach lattice with a lattice norm $\|\cdot\|$ and X^+ be the positive cone of X. We denote by $B(X)$ the unit ball of X, by $S(X)$ the unit sphere of X, and by X^* the dual space of X. We start with auxiliary definitions and results.

Definition 1.1. (See [2,5,20,28].) X is said to be uniformly monotone (UM) if for every $\varepsilon > 0$ there exists $\delta(\varepsilon) > 0$ such that $\|f + g\| > 1 + \delta(\varepsilon)$ whenever $f, g \in X^+$, $\|f\| = 1$ and $\|g\| \geqslant \varepsilon$.

Definition 1.2. (See [5,15,20,28].) X is said to be strictly monotone (STM) if $\|f + g\| > 1$ for all $f, g \in X^+$ with $\|f\| = 1$ and $\|g\| > 0$.

Definition 1.3. (See [22,24].) X is said to be upper (lower) locally uniformly monotone (ULUM, LLUM) if for any $f \in X^+$ with $\|f\| = 1$ and any $\varepsilon > 0$ there is $\delta(f, \varepsilon) > 0$ such that $\|f + g\| > 1 + \delta(f, \varepsilon)$ ($\|f - g\| \leqslant 1 - \delta(f, \varepsilon)$) whenever $g \in X^+$ ($0 \leqslant g \leqslant f$) and $\|g\| \geqslant \varepsilon$.

Definition 1.4. (See [28].) X is said to be weakly uniformly monotone (WUM), if for each positive functional $f^* \in X^*$ and all sequences $f_n, g_n \in X$, $f_n \geqslant g_n \geqslant 0$ with $\|f_n\| = 1$ the condition $\|f_n - g_n\| \to 1$ implies $f^*(g_n) \to 0$.

Definition 1.5. (See [28].) X is said to be weakly uniformly monotone in the second sense (CWUM), if for each positive functional $f^* \in S(X^*)$ and all sequences $f_n, g_n \in X$, $f_n \geqslant g_n \geqslant 0$ with $\|f_n\| = 1$ the condition $f^*(f_n - g_n) \to 1$ implies $\|g_n\| \to 0$.

Localization of the latter two properties leads to the concepts of WULUM (WLLUM) and CWULUM (CWLLUM) spaces, respectively.

Definition 1.6. (See [22].) A Banach lattice X is said to have H^+ property if $\|f - f_n\| \to 0$ whenever $0 \leqslant f_n \leqslant f$ and $f_n \to f$ weakly.

We say that X has the H^+STM property if X has the H^+ property and X is STM.

The following implications are evident:

$$\begin{aligned}
&\mathrm{UM} \Rightarrow \mathrm{ULUM}\ (\mathrm{LLUM}) \Rightarrow \mathrm{WULUM}\ (\mathrm{WLLUM}) \Rightarrow \mathrm{STM},\\
&\mathrm{UM} \Rightarrow \mathrm{WUM} \Rightarrow \mathrm{WULUM}\ (\mathrm{WLLUM}),\\
&\mathrm{UM} \Rightarrow \mathrm{CWUM} \Rightarrow \mathrm{CWLLUM} \Rightarrow \mathrm{H}^+\mathrm{STM} \Rightarrow \mathrm{STM}.
\end{aligned}$$

Let $A(u)$ be an N-function, $\bar{A}(v)$ be the complemented function of $A(u)$, and let $p(u)$ be the right derivative of $A(u)$ [6,26,33,34]. We say that $A(u)$ satisfies Δ_2-condition for large u ($A \in \Delta_2$) if there exist $K > 2$ and $u_0 > 0$ such that $A(2u) \leqslant KA(u)$ for any $u \geqslant u_0$.

Definition 1.7. (See [1].) Let Ω be a bounded and connected open subset of R^n, and let (Ω, Σ, μ) be a nonatomic finite measure space. For any measurable function u on Ω, we define the modular of u by $\rho_A(u) = \int_\Omega A(u(t))\,\mathrm{d}t$. Then the Orlicz–Sobolev space is defined as follows:

$$W_{m,A} = \bigl\{u \in L_A(\Omega)\colon\ \partial^\alpha u \in L_A(\Omega),\ 0 \leqslant |\alpha| \leqslant m\bigr\},$$

where $L_A := L_A(\Omega) = \{u(t)\colon$ there exists $\lambda > 0$, such that $\rho_A(\lambda u) < \infty\}$ is an Orlicz space [6,26,30,33,34], $m \in \{0, 1, 2, \ldots\}$, $\alpha = (\alpha_1, \ldots, \alpha_n)$, $|\alpha| = \alpha_1 + \cdots + \alpha_n$, $\alpha_i (i = 1, \ldots, n)$ are nonnegative integers, $\partial^\alpha u$ is the αth distributional derivative of u.

Let $1 < p < \infty$ be a fixed number. For each $u \in W_{m,A}$, define the norm of u by

$$\|u\|_{m,A} = \Bigl(\sum_{0 \leqslant |\alpha| \leqslant m} \bigl\|\partial^\alpha u\bigr\|^p\Bigr)^{\frac{1}{p}},$$

where $\|u\| = \inf\{k > 0\colon\ \rho_A(u/k) \leqslant 1\}$ is the Luxemburg norm in the Orlicz space L_A. It is well known [1] that $W_{m,A}$ is a Banach space.

Let X be a Banach lattice, $f \in X$ and $K \subset X$ be a nonempty subset. The best approximation operator (called also the projection of X onto K) is defined as follows:

$$P_K(f) = \Bigl\{u \in K\colon\ \|f - u\| = \inf_{h \in K} \|f - h\|\Bigr\}.$$

Let K be a subset of X and $f \in X$. We write $f \geqslant K$ if $f \geqslant g$ for all $g \in K$. Similarly $f \leqslant K$ is defined. For any $f \in X$, a sequence $\{h_n\}$ in K is said to be a minimizing sequence for f if

$$\lim_{n \to \infty} \|f - h_n\| = \inf_{h \in K} \|f - h\| = \operatorname{dist}(f, K).$$

Proposition 1.1. *(See [8].) Let $\gamma = \inf\{t_1\colon (t_1, \dots, t_n) \in \Omega\}$, $\beta = \sup\{t_1\colon (t_1, \dots, t_n) \in \Omega\}$, and let $r \in [\gamma, \beta]$. Assume that $f \in L_A$ satisfies $f(t) \geqslant 0$ for μ-a.e. $t = (t_1, \dots, t_n) \in \Omega$, and $f(t) = 0$ for $t_1 \leqslant r$. Then if $f(t)$ is non-decreasing with respect to t_1, we have*

$$\left\| \int\limits_{E(t)} f(s, t_2, \dots, t_n)\,\mathrm{d}s \right\| \leqslant (\beta - r)\|f\|,$$

where $E(t) = E_r(t_1, \dots, t_n) = \{s \in [r, t_1]\colon (s, t_2, \dots, t_n) \in \Omega\}$.

Proposition 1.2. *(See [28].) Let X be a Banach lattice. The following statements are equivalent*:

(1) *X is STM.*
(2) *For all $f \in X$ and order interval $[a, b] \subset X$ satisfying $f \geqslant [a, b]$ there holds* $\mathrm{Card}(P_{[a,b]}(f)) \leqslant 1$.
(3) *For all $f \in X$ and all sublattices $K \subset X$ satisfying $f \geqslant K$ there holds* $\mathrm{Card}(P_K(f)) \leqslant 1$.

Proposition 1.3. *(See [22].) For any σ-complete Banach lattice X, the following statements are equivalent*:

(1) *X is CWLLUM.*
(2) *X is H^+STM.*
(3) *X is STM and order continuous.*

This paper is devoted to monotonicity properties of Orlicz–Sobolev spaces and their applications to the dominated best approximation problems in the spaces. Other problems such as separability, duality, reflexivity and their comparisons, in Orlicz–Sobolev spaces with parameter, called also Musielak–Orlicz spaces, have been considered by H. Hudzik in [18]. The problems of density of infinitely smooth functions in Orlicz–Sobolev spaces were considered by A. Benkirane and J.P. Gossez in [4] and in Musielak–Orlicz–Sobolev spaces by H. Hudzik in [16]. The problem of embeddings of Orlicz–Sobolev spaces into $C^m(\Omega)$ was investigated in [17]. Various applications of Orlicz–Sobolev spaces to some boundary value problems in differential equations and optimization problems were studied by J.P. Gossez in [11–13].

2. Monotonicity of Orlicz–Sobolev spaces

In this section, we discuss various monotonicity properties of Orlicz–Sobolev spaces. First, we need some agreement on how to define a partial order in the underlying spaces. Clearly, L_A is a Banach lattice in the usual sense, that is, if $x, y \in L_A$ then $x \leqslant y$ means that $x(t) \leqslant y(t)$ for a.e. $t \in \Omega$. But for $m \neq 0$, the Orlicz–Sobolev space $W_{m,A}$ is obviously not a Banach lattice under the same partial order. We shall define a partial order in this space by the following procedure. Let

$$P : W_{m,A} \to \prod_{0 \leqslant |\alpha| \leqslant m} L_A$$

be defined as

$$P(u) = \left(\partial^\alpha u\right)_{0 \leqslant |\alpha| \leqslant m}.$$

Furthermore, define the norm in $\prod_{0 \leqslant |\alpha| \leqslant m} L_A$ as

$$\|x\|_* = \left(\sum_{0 \leqslant |\alpha| \leqslant m} \|x_\alpha\|^p \right)^{\frac{1}{p}}, \quad x = (x_\alpha)_{0 \leqslant |\alpha| \leqslant m} \in \prod_{0 \leqslant |\alpha| \leqslant m} L_A.$$

We equip the product $\prod_{0 \leqslant |\alpha| \leqslant m} L_A$ with the coordinate partial order, that is, for any $x = (x_\alpha)_{0 \leqslant |\alpha| \leqslant m}$, $y = (y_\alpha)_{0 \leqslant |\alpha| \leqslant m} \in \prod_{0 \leqslant |\alpha| \leqslant m} L_A$,

$$x \vee y = (x_\alpha \vee y_\alpha)_{0 \leqslant |\alpha| \leqslant m},$$
$$x \leqslant y \quad \text{iff} \quad x_\alpha \leqslant y_\alpha \quad \text{for all } \alpha \text{ with } 0 \leqslant |\alpha| \leqslant m.$$

Then $\prod_{0\leqslant|\alpha|\leqslant m} L_A$ is a Banach lattice and $P: W_{m,A} \to \prod_{0\leqslant|\alpha|\leqslant m} L_A$ is an isometric linear operator, which induces the partial order in $W_{m,A}$ from the product $\prod_{0\leqslant|\alpha|\leqslant m} L_A$. In other words we write for $u, v \in W_{m,A}$ that

$$u \leqslant v \quad \text{whenever} \quad \partial^\alpha u(t) \leqslant \partial^\alpha v(t) \quad \text{for all } \alpha = (\alpha_1, \ldots, \alpha_n) \text{ with } 0 \leqslant |\alpha| \leqslant m, \text{ a.e. in } \Omega.$$

In this sense, we may consider $W_{m,A}$ as a closed subspace of $\prod_{0\leqslant|\alpha|\leqslant m} L_A$. Further we shall identify $x \in W_{m,A}$ with $P(x)$, that is with $(\partial^\alpha x)_{0\leqslant|\alpha|\leqslant m}$. Thus for $x \in W_{m,A}$, the element $|x|$ is identified with $(|\partial^\alpha x|)_{0\leqslant|\alpha|\leqslant m}$. Notice that $|x|$ does not need to belong to $W_{m,A}$ even when x is in the space. So in general $W_{m,A}$ is not a sublattice of $\prod_{0\leqslant|\alpha|\leqslant m} L_A$.

We shall study the monotonicity properties of the lattice $W_{m,A}$ equipped with the partial order defined above.

Lemma 2.1. *Let X be a Banach lattice. X is STM if and only if for all $x, y \in X$ with $y \geqslant 0$ and $y \neq 0$, the conditions $\|x + \lambda y\| = \||x| + \lambda y\|$ and $|x + \lambda y| > |x|$ for all $\lambda \in (0, 1]$ imply $\|x + y\| > \|x\|$.*

Proof. Sufficiency. For any $x, y \in X^+$ with $y \neq 0$, we have

$$\|x + \lambda y\| = \big\||x| + \lambda y\big\| \quad \text{and} \quad |x + \lambda y| > |x| \quad \text{for all } \lambda \in (0, 1].$$

By the assumption, we have $\|x + y\| > \|x\|$, i.e., X is STM.

Necessity. For any $x, y \in X$ with $0 \leqslant y \neq 0$, if $\|x + \lambda y\| = \||x| + \lambda y\|$ and $|x + \lambda y| > |x|$ for all $\lambda \in (0, 1]$, then $\|x + y\| = \||x| + y\|$ and since $|x| + y > |x|$ and X is STM, we have $\|x + y\| = \||x| + y\| > \|x\|$. □

Theorem 2.1. *The following statements are equivalent*:

(1) $\prod_{0\leqslant|\alpha|\leqslant m} L_A$ *is UM.*
(2) $W_{m,A}$ *is UM.*
(3) $W_{m,A}$ *is ULUM* (*LLUM*).
(4) $W_{m,A}$ *is STM.*
(5) $A \in \Delta_2$.

Proof. The implications (1) ⇒ (2) ⇒ (3) ⇒ (4) are clear.

(5) ⇒ (1). Since $A \in \Delta_2$, from [22,28,32], we know that L_A is UM. It is easy to deduce that $\prod_{0\leqslant|\alpha|\leqslant m} L_A$ is UM.

(4) ⇒ (5). Take $\gamma = \inf\{t_1\colon (t_1, \ldots, t_n) \in \Omega\}$, $\beta = \sup\{t_1\colon (t_1, \ldots, t_n) \in \Omega\}$, and $\delta = A^{-1}(\frac{1}{2\mu\Omega})$. If $A \notin \Delta_2$, there exists a nonnegative sequence $\{u_i\} \uparrow \infty$, satisfying

$$A\left(\left(1 + \frac{1}{i}\right)u_i\right) > 2^i A(u_i), \quad u_i \geqslant (i+2)\delta, \qquad u_1 \geqslant A^{-1}\left(\frac{1}{\mu\Omega}\right) \quad (i \in \mathbb{N}).$$

For any $i \in \mathbb{N}$, let $\lambda_i = \frac{1}{2^{i+1}A(u_i)}$. Then $0 \leqslant \lambda_i \downarrow 0$, and $\sum_{i=1}^\infty \lambda_i \leqslant \sum_{i=1}^\infty \frac{1}{2^{i+1}A(u_1)} \leqslant \sum_{i=1}^\infty \frac{1}{2^{i+1}}\mu\Omega = \frac{1}{2}\mu\Omega$. Take $\delta_1 \in (\gamma, \beta)$ such that setting $E = \{(t_1, \ldots, t_n) \in \Omega\colon t_1 > \delta_1\}$, we have $\mu E = \sum_{i=1}^\infty \lambda_i$. Then choose $\delta_2 \in (\delta_1, \beta]$ and $E_1 = \{(t_1, \ldots, t_n) \in \Omega\colon \delta_1 < t_1 \leqslant \delta_2\}$ such that $\mu E_1 = \lambda_1$. By the induction process we find $\delta_{k+1} \in (\delta_k, \beta]$ and $E_k = \{(t_1, \ldots, t_n) \in \Omega\colon \delta_k < t_1 \leqslant \delta_{k+1}\}$ such that $\mu E_k = \lambda_k$, $k = 2, 3, \ldots$. It is obvious that $A(u_i)\mu E_i = \frac{1}{2^{i+1}}$, $(i \in \mathbb{N})$. Choose also for every $k \in \mathbb{N}$, $\beta_k \in (\delta_k, \delta_{k+1})$ and $F_k = \{(t_1, \ldots, t_n) \in E_k\colon t_1 > \beta_k\}$ satisfying $\mu F_k = \frac{\lambda_k}{2}$.

Define now $y(t) = \delta$ on Ω and

$$x(t) = \begin{cases} 0, & t \in \Omega \setminus E, \\ u_k - \delta, & t \in F_k,\ k \in \mathbb{N}, \end{cases}$$

and define $x(t)$ on $E_k \setminus F_k$ in such a way that $x(t) \geqslant 0$ a.e., $x(t)$ is infinitely differentiable on Ω, non-decreasing with respect to t_1, and is constant with respect to $t_2, \ldots, t_n$. Then

$$\begin{aligned}\rho_A(x + y) &= \int_E A\big(x(t) + \delta\big)\,dt + \int_{\Omega\setminus E} A(\delta)\,dt \\ &= \sum_{i=1}^\infty \left[\int_{F_i} A(u_i)\,dt + \int_{E_i\setminus F_i} A\big(x(t) + \delta\big)\,dt\right] + \int_{\Omega\setminus E} A(\delta)\,dt\end{aligned}$$

$$\leqslant \sum_{i=1}^{\infty}\Big[\int_{F_i} A(u_i)\,\mathrm{d}t + \int_{E_i\setminus F_i} A(u_i)\,\mathrm{d}t\Big] + \int_{\Omega\setminus E} A(\delta)\,\mathrm{d}t$$
$$= \sum_{i=1}^{\infty} A(u_i)\mu E_i + A(\delta)\mu\Omega = \sum_{i=1}^{\infty}\frac{1}{2^{i+1}} + \frac{1}{2} = 1,$$

which implies that $\|x\| \leqslant \|x+y\| \leqslant 1$. On the other hand, given any $\varepsilon > 0$ we take $i_0 > \frac{2}{\varepsilon}$, then by $u_i > (i+2)\delta$, we have for $i \geqslant i_0$,

$$(1+\varepsilon)(u_i-\delta) > \left(1+\frac{2}{i_0}\right)(u_i-\delta) > \left(1+\frac{2}{i}\right)(u_i-\delta) = \left(1+\frac{1}{i}\right)u_i + \left(\frac{1}{i}u_i - \Big(1+\frac{2}{i}\Big)\delta\right) > \left(1+\frac{1}{i}\right)u_i.$$

Thus we obtain that

$$\rho_A\big((1+\varepsilon)x\big) > \sum_{i=i_0}^{\infty}\int_{F_i} A\big((1+\varepsilon)x(t)\big)\,\mathrm{d}t = \sum_{i=i_0}^{\infty}\int_{F_i} A\big((1+\varepsilon)(u_i-\delta)\big)\,\mathrm{d}t$$
$$= \sum_{i=i_0}^{\infty} A\big((1+\varepsilon)(u_i-\delta)\big)\mu F_i = \sum_{i=i_0}^{\infty}\frac{\lambda_i}{2} A\big((1+\varepsilon)(u_i-\delta)\big)$$
$$> \sum_{i=i_0}^{\infty}\frac{\lambda_i}{2} A\left(\left(1+\frac{1}{i}\right)u_i\right) \geqslant \sum_{i=i_0}^{\infty}\frac{\lambda_i}{2}2^i A(u_i) = \sum_{i=i_0}^{\infty}\frac{1}{4} = \infty,$$

which yields that $\|(1+\varepsilon)x\| \geqslant 1$. By the arbitrariness of $\varepsilon > 0$ we get $1 \leqslant \|x\| \leqslant \|x+y\| \leqslant 1$, i.e., $\|x\| = \|x+y\| = 1$. Notice that by the construction of x and y, they are both nonnegative in the sense of the partial order in $W_{m,A}$, that is, $\partial^\alpha x(t) \geqslant 0$, $\partial^\alpha y(t) \geqslant 0$ a.e. in Ω for any $|\alpha| \leqslant m$. Hence

$$\|x+y\|_{m,A} = \left(\sum_{0\leqslant|\alpha|\leqslant m}\big\|\partial^\alpha(x+y)\big\|^p\right)^{\frac{1}{p}} = \left(\|x+y\|^p + \sum_{1\leqslant|\alpha|\leqslant m}\big\|\partial^\alpha(x+y)\big\|^p\right)^{\frac{1}{p}}$$
$$= \left(\|x\|^p + \sum_{1\leqslant|\alpha|\leqslant m}\big\|\partial^\alpha x\big\|^p\right)^{\frac{1}{p}} = \left(\sum_{0\leqslant|\alpha|\leqslant m}\big\|\partial^\alpha x\big\|^p\right)^{\frac{1}{p}} = \|x\|_{m,A},$$

which shows that $W_{m,A}$ does not have the STM property. We wish to point out that this implication is related to the observation made in [15]. □

In [7], Shutao Chen and Changying Hu proved that if $W_{m,A}$ is rotund, then A is strictly convex. Thus, by Theorem 2.1, we obtain the following result.

Theorem 2.2. *The following conditions are equivalent*:

(1) $W_{m,A}$ *is locally uniformly rotund.*
(2) $W_{m,A}$ *is rotund.*
(3) $A \in \Delta_2$ *and* A *is strictly convex.*

Proof. The implication (1) ⇒ (2) is obvious.

(2) ⇒ (3). From Theorem 2.1 and [7], we only need to prove that rotundity implies strict monotonicity. Indeed, for any $x, y \in W_{m,A}$ with $x \geqslant y \geqslant 0$, $y \neq 0$, $x \neq y$ and $\|x\|_{m,A} = 1$, we have $y \leqslant \frac{x+y}{2} \leqslant x$. By the rotundity of $W_{m,A}$, we get $\|y\|_{m,A} \leqslant \|\frac{x+y}{2}\|_{m,A} < 1 = \|x\|_{m,A}$, which means that $W_{m,A}$ is strictly monotone.

(3) ⇒ (1). If $A \in \Delta_2$ and A is strictly convex, then from [6] we know that L_A is locally uniformly rotund. Then it is easy to prove that $\prod_{0\leqslant|\alpha|\leqslant m} L_A$ is locally uniformly rotund. So $W_{m,A}$ is locally uniformly rotund. □

3. Monotonicity and best approximation

In [22,28], the best approximation problem in K for $f \in X$ was discussed for the case $f - K \geqslant 0$, where $f \in X$ and K was a sublattice of X. Instead of those restrictions, we will consider more general case, that is, we only require that K is a convex set and $K - f$ is so-called absolutely direct set.

Let $K \subset W_{m,A}$. We say that K is an absolutely direct set if for any $x, y \in K$ there exists $z \in K$ such that

$$|z| \leqslant |x| \wedge |y|, \quad \text{that is,} \quad \left|\partial^\alpha z(t)\right| \leqslant \min\left\{\left|\partial^\alpha x(t)\right|, \left|\partial^\alpha y(t)\right|\right\}$$

for all α with $0 \leqslant |\alpha| \leqslant m$ and for a.e. $t \in \Omega$.

For any fixed subset K of $W_{m,A}$, we introduce the set

$$D(K) = \{f \in W_{m,A}\colon\ K - f \text{ is an absolutely direct set}\}.$$

Now, we turn to the best approximation problem in Orlicz–Sobolev spaces. Let us first agree on the following notation. If $x \in W_{m,A}$, then

$$\|x\|_* := \|x\|_{m,A} = \left(\sum_{0 \leqslant |\alpha| \leqslant m} \left\|\partial^\alpha x\right\|^p\right)^{\frac{1}{p}}.$$

But when $x = (x_\alpha)_{0 \leqslant \alpha \leqslant m} \in \prod_{0 \leqslant |\alpha| \leqslant m} L_A$ and we are not sure that $x \in W_{m,A}$ (that is, we do not know that there exists $y \in W_{m,A}$ such that $P(y) = x$), then by $\|x\|_*$ we mean the symbol defined earlier, that is,

$$\|x\|_* := \left(\sum_{0 \leqslant |\alpha| \leqslant m} \|x_\alpha\|^p\right)^{\frac{1}{p}}.$$

Theorem 3.1 *(Uniqueness). The following statements are equivalent*:

(1) *For any convex subset K of $W_{m,A}$ and any $f \in D(K)$ there holds* $\operatorname{Card}(P_K(f)) \leqslant 1$.
(2) *For any closed convex subset K of $W_{m,A}$ and any $f \in D(K)$ there holds* $\operatorname{Card}(P_K(f)) \leqslant 1$.
(3) *$W_{m,A}$ is STM.*
(4) $A \in \Delta_2$.
(5) *$\prod_{0 \leqslant |\alpha| \leqslant m} L_A$ is STM.*

Proof. The implication (1) $\Rightarrow$ (2) is obvious.

(2) $\Rightarrow$ (3). For a contrary assume (2) is satisfied but (3) is not. Then by Lemma 2.1, there exist $x, y \in W_{m,A}$ such that $y \geqslant 0$, $y \neq 0$, $\|x + \lambda y\|_{m,A} = \||x| + \lambda y\|_*$, $|x + \lambda y| > |x|$ for all $\lambda \in (0, 1]$, and $\|x + y\|_{m,A} = \|x\|_{m,A}$.

Define $K = \{\lambda y\colon\ \lambda \in [0, 1]\}$. Then K is a closed convex subset of $W_{m,A}$ and $|x + \lambda y| \geqslant |x|$ for any $\lambda \in [0, 1]$, so $-x \in D(K)$. Moreover, for all $\lambda y \in K$,

$$\|x\|_{m,A} = \|x + y\|_{m,A} = \left\||x| + y\right\|_* \geqslant \left\||x| + \lambda y\right\|_* \geqslant \|x + \lambda y\|_{m,A} \geqslant \|x\|_{m,A},$$

which yields $\|x + \lambda y\|_{m,A} = \||x| + \lambda y\|_* = \|x\|_{m,A}$ for all $\lambda \in [0, 1]$. This shows that $K = P_K(-x)$, and it is a contradiction.

(3) $\Rightarrow$ (4). See Theorem 2.1.

(4) $\Rightarrow$ (5). From [22,28,32], we know that if $A \in \Delta_2$, then L_A is STM. Thus for any $x = (x_\alpha)$, $y = (y_\alpha) \in \prod_{0 \leqslant |\alpha| \leqslant m} L_A$ with $\|x\|_* = 1$ and $\|y\|_* > 0$, we have

$$\|x + y\|_* = \left(\sum_{0 \leqslant |\alpha| \leqslant m} \|x_\alpha + y_\alpha\|^p\right)^{\frac{1}{p}} = \left(\sum_{0 \leqslant |\alpha| \leqslant m} \|x_\alpha\|^p \left(\left\|\frac{x_\alpha}{\|x_\alpha\|} + \frac{y_\alpha}{\|x_\alpha\|}\right\|^p\right)\right)^{\frac{1}{p}} > \left(\sum_{0 \leqslant |\alpha| \leqslant m} \|x_\alpha\|^p\right)^{\frac{1}{p}} = 1.$$

This shows that $\prod_{0 \leqslant |\alpha| \leqslant m} L_A$ is STM, whence $W_{m,A}$ is STM too.

(5) $\Rightarrow$ (1). For any subset K of $W_{m,A}$ and any $f \in D(K)$, pick up any $x, y \in P_K(f)$. Then

$$\|x - f\|_{m,A} = \|y - f\|_{m,A} = \operatorname{dist}(f, K) = d.$$

Set $u = x - f$, $v = y - f$. Since $f \in D(K)$, there exists $z \in K - f$ such that $|z| \leqslant |u| \wedge |v|$. Observe that $|z| \leqslant |u|$ and $d \leqslant \|z\|_* \leqslant \|u\|_* = d$. It follows from STM of $\prod_{0 \leqslant |\alpha| \leqslant m} L_A$ that $|z| = |u|$, that is, $|\partial^\alpha z| = |\partial^\alpha u|$ for every $0 \leqslant |\alpha| \leqslant m$. Similarly we have $|z| = |v|$, and so

$$|u| = |v| \quad \text{in} \prod_{0 \leqslant |\alpha| \leqslant m} L_A.$$

On the other hand, since K is convex, we have $\frac{u+v}{2} \in K - f$. Therefore,

$$d \leqslant \left\| \frac{u+v}{2} \right\|_{m,A} \leqslant \frac{\|u\|_{m,A} + \|v\|_{m,A}}{2} = d.$$

But $|\frac{u+v}{2}| \leqslant \frac{|u|+|v|}{2} = |u| = |v|$. It follows from STM of $\prod_{0 \leqslant |\alpha| \leqslant m} L_A$ again that

$$\left| \frac{u+v}{2} \right| = |u| = |v| = |u| \vee |v| \quad \text{in} \prod_{0 \leqslant |\alpha| \leqslant m} L_A.$$

Recalling that $\prod_{0 \leqslant |\alpha| \leqslant m} L_A$ being a Banach lattice has the property $|u+v| + |u-v| = 2(|u| \vee |v|)$, we get $|u - v| = 0$, i.e., $x = y$. This shows that $\mathrm{Card}(P_K(f)) \leqslant 1$. $\square$

Theorem 3.2 *(Existence). The space $W_{m,A}$ has the property that for any convex set K in $W_{m,A}$ and any $f \in D(K)$, $K - f$ has a minimizing Cauchy sequence if and only if $A \in \Delta_2$.*

Proof. Necessity. If A does not satisfy Δ_2-condition, then there exists a nonnegative sequence $\{u_k\} \uparrow \infty$, satisfying

$$A\left(\left(1 + \frac{1}{k}\right)u_k\right) > 2^k A(u_k) \quad (k \in \mathbb{N}).$$

For any $k \in \mathbb{N}$, let $\lambda_k = \frac{1}{2^k A(u_k)}$. Take $\gamma = \inf\{t_1\colon (t_1, \ldots, t_n) \in \Omega\}$ and $\beta = \sup\{t_1\colon (t_1, \ldots, t_n) \in \Omega\}$. Since $\lambda_k < \frac{1}{2^k}$ for k large enough, we have that $\sum_{k=1}^\infty \lambda_k$ is convergent. Without loss of generality, we may assume that $\sum_{k=1}^\infty \lambda_k < \mu\Omega$. Let $\delta_0 \in (\gamma, \beta)$ be such that $G_0 = \{t \in \Omega\colon t_1 \leqslant \delta_0\}$ satisfies $\mu G_0 = \mu\Omega - \sum_{k=1}^\infty \lambda_k$. Take $\delta_1 \in (\delta_0, \beta]$ and $G_1 = \{t \in \Omega\colon \delta_0 < t_1 \leqslant \delta_1\}$ such that $\mu G_1 = \lambda_1$. By induction, let for every $k \in \mathbb{N}$, $\delta_k \in (\delta_{k-1}, \beta]$, $G_k = \{t \in \Omega\colon \delta_{k-1} < t_1 \leqslant \delta_k\}$ be such that $\mu G_k = \lambda_k$. It is obvious that $A(u_k)\mu G_k = \frac{1}{2^k}$ for all $k \in \mathbb{N}$.

Define $x_j(t) = \sum_{k=j+1}^\infty u_k \chi_{G_k}(t)$. Then $0 \leqslant x_j \downarrow 0$ and

$$\rho_A(x_j) = \sum_{k=j+1}^\infty A(u_k)\mu G_k = \sum_{k=j+1}^\infty \frac{1}{2^k} = \frac{1}{2^j} < 1,$$

$$\begin{aligned}\rho_A\left(\left(1 + \frac{1}{j}\right)x_j\right) &= \sum_{k=j+1}^\infty A\left(\left(1 + \frac{1}{j}\right)u_k\right)\mu G_k \geqslant \sum_{k=j+1}^\infty A\left(\left(1 + \frac{1}{k}\right)u_k\right)\mu G_k \\ &> \sum_{k=j+1}^\infty 2^k A(u_k)\mu G_k = \sum_{k=j+1}^\infty 1 = \infty.\end{aligned}$$

Therefore, $1 > \|x_j\| \geqslant \frac{1}{1+(1/j)} \to 1$, whence $\|x_j\| \to 1$ as $j \to \infty$.

Let $f_j(t) = f_j(t_1, \ldots, t_n)$ be a function obtained by integrating m times the function $x_j(s) = x_j(s_1, \ldots, s_n)$ with respect to the first variable s_1 over the interval (δ_j, t_1), that is,

$$f_j(t) = \overbrace{\left(\int_{\delta_j}^{t_1} \cdots \left(\int_{\delta_j}^{t_1} x_j(u, t_2, \ldots, t_n)\,\mathrm{d}u\right) \cdots \mathrm{d}u\right)}^{m}.$$

Then $0 \leqslant f_j \downarrow 0$ a.e. in Ω,

$$\partial_1^m f_j = x_j \quad \text{and} \quad \partial_i f_j = 0 \quad \text{for } i \neq 1.$$

For any α $(0 \leqslant |\alpha| \leqslant m-1)$, since $x_j(t)$ is non-decreasing with respect to t_1, by Proposition 1.1, we have

$$\|\partial_1^\alpha f_j\| = \Bigg\| \overbrace{\Bigg(\int_{\delta_j}^{t_1} \cdots \Bigg(\int_{\delta_j}^{t_1}}^{m-|\alpha|} x_j(u, t_2, \ldots, t_n)\,\mathrm{d}u\Bigg) \cdots \mathrm{d}u\Bigg)\Bigg\| \leqslant (\beta - \delta_j)^{m-|\alpha|}\|x_j\| \to 0 \quad \text{as } j \to \infty.$$

So $\lim_{j\to\infty} \|f_j\|_{m,A} = \lim_{j\to\infty} \|x_j\| = 1$. Set

$$K = \operatorname{conv}\{f_j\}_{j=1}^\infty = \Bigg\{ \sum_{i=1}^j \lambda_i f_i \colon 0 \leqslant \lambda_i \leqslant 1,\ \sum_{i=1}^j \lambda_i = 1 \text{ for any } j \Bigg\}.$$

Notice that $\partial_1^\alpha f_i(t) \geqslant \partial_1^\alpha f_j(t)$ a.e. in Ω for $i = 1, \ldots, j$ and $0 \leqslant |\alpha| \leqslant m$. Hence for any $\sum_{i=1}^j \lambda_i f_i \in K$, $\sum_{i=1}^j \lambda_i f_i \geqslant f_j \geqslant 0$ in the sense of order in $W_{m,A}$. Thus $0 \in D(K)$, that is, K is absolutely direct set and $\|\sum_{i=1}^j \lambda_i f_i\|_{m,A} \geqslant \|f_j\|_{m,A}$. It follows that $\{f_j\}$ is a minimizing sequence in K for zero. The sequence $\{f_j\}$ is not Cauchy. In fact, in view of $\mu(G_k) = 1/2^k A(u_k)$, for all $i, j \in \mathbb{N}$, $j > i$ we have

$$\rho_A\big(2(x_j - x_i)\big) = \sum_{k=i+1}^j A(2u_k)\mu(G_k) > \sum_{k=i+1}^j 2^k A(u_k)\mu(G_k) = j - i \geqslant 1.$$

Hence for all $j > i$,

$$\|f_j - f_i\|_{m,A} \geqslant \|x_j - x_i\| \geqslant 1/2,$$

which shows our claim.

Sufficiency follows from the next Theorem 3.3. □

Theorem 3.3 *(Stability). For any convex absolutely direct set K of $W_{m,A}$, each minimizing sequence $\{x_j\}$ of K is a Cauchy sequence in $W_{m,A}$ if and only if $A \in \Delta_2$.*

Proof. *Necessity* follows from the proof of necessity in Theorem 3.2. In fact under the assumption that A does not satisfy the Δ_2-condition we found a convex absolutely direct set K and a minimizing sequence $\{f_j\}$ which is not Cauchy.

Sufficiency. Assume $A \in \Delta_2$. Then $\prod_{0\leqslant|\alpha|\leqslant m} L_A$ is UM by Theorem 2.1.

For any absolutely direct convex set $K \subset W_{m,a}$, let $\{x_j\}$ be a minimizing sequence of K for zero. Then $\lim_{j\to\infty} \|x_j\|_{m,A} = \inf_{x\in K} \|x\|_{m,A} = d(K, 0) := d$.

If $\{x_j\}$ is not a Cauchy sequence, there are subsequences $\{x_{n_k}\}$, $\{x_{m_k}\}$ of $\{x_j\}$ and $\varepsilon_0 > 0$ such that $\|x_{n_k} - x_{m_k}\|_{m,A} \geqslant \varepsilon_0$. Since K is an absolutely direct set, there exist $z_k \in K$ such that

$$|z_k| \leqslant |x_{n_k}| \wedge |x_{m_k}| \quad \text{in} \prod_{0\leqslant|\alpha|\leqslant m} L_A. \tag{3.1}$$

By (3.1), we have $d \leqslant \|z_k\|_{m,A} \leqslant \|x_{n_k}\|_{m,A} \to d$, i.e., $\|x_{n_k}\|_{m,A} - \|z_k\|_{m,A} \to 0$. Since $|z_k| \leqslant |x_{n_k}|$, by the UM property of $\prod_{0\leqslant|\alpha|\leqslant m} L_A$, we get

$$\big\| |x_{n_k}| - |z_k| \big\|_* \to 0. \tag{3.2}$$

Likewise, $\||x_{m_k}| - |z_k|\|_* \to 0$. For any α with $0 \leqslant |\alpha| \leqslant m$, define the sets

$$\Omega_{n,k}^{\alpha,0} = \big\{t \in \Omega \colon \big|\partial^\alpha x_{n_k}(t)\big| + \big|\partial^\alpha z_k(t)\big| = \big|\partial^\alpha x_{n_k}(t) + \partial^\alpha z_k(t)\big|\big\}, \qquad \Omega_{n,k}^{\alpha,1} = \Omega \setminus \Omega_{n,k}^{\alpha,0}.$$

Then by (3.2) we have for any $0 \leqslant |\alpha| \leqslant m$,

$$0 \leftarrow \big\| |\partial^\alpha x_{n_k}| - |\partial^\alpha z_k| \big\| = \big\| \big|\big(\partial^\alpha x_{n_k} - \partial^\alpha z_k\big)\chi_{\Omega_{n,k}^{\alpha,0}}\big| + \big|\big(\partial^\alpha x_{n_k} + \partial^\alpha z_k\big)\chi_{\Omega_{n,k}^{\alpha,1}}\big| \big\|,$$

which implies that

$$\big\|\big(\partial^\alpha x_{n_k} - \partial^\alpha z_k\big)\chi_{\Omega_{n,k}^{\alpha,0}}\big\| \to 0, \qquad \big\|\big(\partial^\alpha x_{n_k} + \partial^\alpha z_k\big)\chi_{\Omega_{n,k}^{\alpha,1}}\big\| \to 0. \tag{3.3}$$

Since K is convex, $\frac{x_{n_k}+z_k}{2} \in K$ and $|\frac{x_{n_k}+z_k}{2}| \leqslant \frac{|x_{n_k}|+|z_k|}{2} \leqslant |x_{n_k}|$. Thus $d \leqslant \|\frac{x_{n_k}+z_k}{2}\|_{m,A} \leqslant \|x_{n_k}\|_{m,A} \to d$. By the UM property of $\prod_{0\leqslant|\alpha|\leqslant m} L_A$, we get

$$\left\| |x_{n_k}| - \left|\frac{x_{n_k}+z_k}{2}\right| \right\|_* \to 0.$$

Consequently, for any α with $0 \leqslant |\alpha| \leqslant m$, we have

$$\left\| \left|\partial^\alpha x_{n_k}\right| - \left|\frac{\partial^\alpha x_{n_k}+\partial^\alpha z_k}{2}\right| \right\| \to 0, \qquad \left\| \left|(\partial^\alpha x_{n_k})\chi_{\Omega_{n,k}^{\alpha,1}}\right| - \left|\left(\frac{\partial^\alpha x_{n_k}+\partial^\alpha z_k}{2}\right)\chi_{\Omega_{n,k}^{\alpha,1}}\right| \right\| \to 0. \tag{3.4}$$

Therefore, by (3.3) and (3.4),

$$\left\|\partial^\alpha x_{n_k}\chi_{\Omega_{n,k}^{\alpha,1}}\right\| \leqslant \left\| \left(\left|\partial^\alpha x_{n_k}\chi_{\Omega_{n,k}^{\alpha,1}}\right| - \left|\left(\frac{\partial^\alpha x_{n_k}+\partial^\alpha z_k}{2}\right)\chi_{\Omega_{n,k}^{\alpha,1}}\right| \right) \right\| + \left\| \left(\frac{\partial^\alpha x_{n_k}+\partial^\alpha z_k}{2}\right)\chi_{\Omega_{n,k}^{\alpha,1}} \right\| \to 0. \tag{3.5}$$

Hence, by (3.3) and (3.5) for any α with $0 \leqslant |\alpha| \leqslant m$ we have

$$\begin{aligned}
\left\|\partial^\alpha x_{n_k} - \partial^\alpha z_k\right\| &= \left\|\left(\partial^\alpha x_{n_k} - \partial^\alpha z_k\right)\chi_{\Omega_{n,k}^{\alpha,0}} + \left(\partial^\alpha x_{n_k} - \partial^\alpha z_k\right)\chi_{\Omega_{n,k}^{\alpha,1}}\right\| \\
&\leqslant \left\|\left(\partial^\alpha x_{n_k} - \partial^\alpha z_k\right)\chi_{\Omega_{n,k}^{\alpha,0}}\right\| + \left\|\left|\partial^\alpha x_{n_k}\chi_{\Omega_{n,k}^{\alpha,1}}\right| + \left|\partial^\alpha z_k\chi_{\Omega_{n,k}^{\alpha,1}}\right|\right\| \\
&\leqslant \left\|\left(\partial^\alpha x_{n_k} - \partial^\alpha z_k\right)\chi_{\Omega_{n,k}^{\alpha,0}}\right\| + \left\|2\partial^\alpha x_{n_k}\chi_{\Omega_{n,k}^{\alpha,1}}\right\| \to 0.
\end{aligned}$$

Then $\|x_{n_k} - z_k\|_{m,A} \to 0$. Likewise, $\|x_{m_k} - z_k\|_{m,A} \to 0$. Hence $\|x_{n_k} - x_{m_k}\|_{m,A} \to 0$.

This contradicts the assumption that $\|x_{n_k} - x_{m_k}\|_{m,A} \geqslant \varepsilon_0$. □

Remark 3.1. It is proved in [21] that for any nonempty convex, closed and bounded set K in a Banach space X, each minimizing sequence $\{x_j\}$ of K has a Cauchy subsequence if and only if X is reflexive. Theorem 3.3 shows that we have even a little more for convex absolutely direct sets K in $W_{m,A}$ without reflexivity of $W_{m,A}$ but for smaller class of sets K.

Theorem 3.4 *(Continuity of best approximation operator). If $A \in \Delta_2$, then for any closed convex subset K of $W_{m,A}$ and $f \in D(K)$,* $\mathrm{Card}(P_K(f)) = 1$ *and the best approximation operator $P_K : D(K) \to K$ is continuous.*

Proof. (1) If $A \in \Delta_2$, by Theorems 3.1 and 3.2 we know that $\mathrm{Card}(P_K(f)) \leqslant 1$ and $K - f$ has a minimizing Cauchy sequence. Since K is closed and $W_{m,A}$ is STM, we get $\mathrm{Card}(P_K(f)) = 1$.

(2) Let $f_j \in D(K)$, $P_K(f_j) = \{x_j\}$ $(j = 0, 1, 2, \ldots)$. If $\|f_j - f_0\|_{m,A} \to 0$, we will prove that $\|x_j - x_0\|_{m,A} \to 0$. First, by Theorem 2.1, $A \in \Delta_2$ implies that $\prod_{0\leqslant|\alpha|\leqslant m} L_A$ is UM.

Since $f_j \in D(K)$ and K is convex, there exist $y_j \in K - f_j$ such that

$$|y_j| \leqslant |x_0 - f_j| \wedge \left|\frac{x_j + x_0}{2} - f_j\right| \wedge |x_j - f_j|.$$

Then $\|y_j\|_{m,A} \leqslant \|x_j - f_j\|_{m,A}$, which yields $y_j + f_j \in P_K(f_j)$, and by the first part $y_j = x_j - f_j$. Then $|x_j - f_j| \leqslant |x_0 - f_j| \wedge |\frac{x_j+x_0}{2} - f_j|$.

Set $d_j = \mathrm{dist}(f_j, K) = \|x_j - f_j\|_{m,A}$ $(j = 0, 1, 2, \ldots)$. Then $\|f_j - f_0\|_{m,A} \to 0$ implies that $d_j \to d_0$. Moreover, since $0 \leqslant \|x_0 - f_j\|_{m,A} - \|x_0 - f_0\|_{m,A} \leqslant \|f_j - f_0\|_{m,A} \to 0$, we have

$$d_0 = \lim_{j\to\infty} d_j = \lim_{j\to\infty} \|x_j - f_j\|_{m,A} \leqslant \lim_{j\to\infty} \|x_0 - f_j\|_{m,A} = \|x_0 - f_0\|_{m,A} = d_0,$$

i.e., $\lim_{n\to\infty} \|x_j - f_j\|_{m,A} = \lim_{j\to\infty} \|x_0 - f_j\|_{m,A} = \|x_0 - f_0\|_{m,A}$. But $|x_j - f_j| \leqslant |x_0 - f_j|$, by the UM property of $\prod_{0\leqslant|\alpha|\leqslant m} L_A$, we have

$$\left\| |x_j - f_j| - |x_0 - f_j| \right\|_* \to 0. \tag{3.6}$$

For any $j \in \mathbb{N}$ and any α with $0 \leqslant |\alpha| \leqslant m$, set

$$\Omega_j^{\alpha,0} = \left\{ t \in \Omega\colon \left|\partial^\alpha x_j(t) - \partial^\alpha f_j(t)\right| + \left|\partial^\alpha x_0(t) - \partial^\alpha f_j(t)\right| = \left|\partial^\alpha x_j(t) - \partial^\alpha x_0(t)\right| \right\},$$

and $\Omega_j^{\alpha,1} = \Omega \setminus \Omega_j^{\alpha,0}$. Then $\Omega = \Omega_j^{\alpha,0} \cup \Omega_j^{\alpha,1}$ and

$$\big\| \big|\partial^\alpha x_j - \partial^\alpha f_j\big| - \big|\partial^\alpha x_0 - \partial^\alpha f_j\big| \big\| = \big\| \big|\big(\partial^\alpha x_j - \partial^\alpha x_0\big)\chi_{\Omega_n^{\alpha,1}}\big| + \big|\big(\partial^\alpha x_j + \partial^\alpha x_0 - 2\partial^\alpha f_j\big)\chi_{\Omega_j^{\alpha,0}}\big| \big\|. \tag{3.7}$$

Thus by (3.6),

$$\big\|\big(\partial^\alpha x_j - \partial^\alpha x_0\big)\chi_{\Omega_j^{\alpha,1}}\big\| \to 0, \qquad \big\|\big(\partial^\alpha x_j + \partial^\alpha x_0 - 2\partial^\alpha f_j\big)\chi_{\Omega_j^{\alpha,0}}\big\| \to 0. \tag{3.8}$$

Since $|x_j - f_j| \leqslant |\frac{x_j+x_0}{2} - f_j|$ and

$$\begin{aligned} d_0 &= \lim_{j\to\infty} \|x_j - f_j\|_{m,A} \leqslant \lim_{j\to\infty} \left\| \frac{x_j + x_0}{2} - f_j \right\|_{m,A} \leqslant \lim_{j\to\infty} \left(\frac{\|x_j - f_j\|_{m,A}}{2} + \frac{\|x_0 - f_j\|_{m,A}}{2} \right) \\ &= \|x_0 - f_0\|_{m,A} = d_0, \end{aligned}$$

the UM property of $\prod_{0\leqslant|\alpha|\leqslant m} L_A$ implies $\||\frac{x_j+x_0}{2} - f_j| - |x_j - f_j|\|_* \to 0$. Consequently, $\||\frac{\partial^\alpha x_j+\partial^\alpha x_0}{2} - \partial^\alpha f_j| - |\partial^\alpha x_j - \partial^\alpha f_j|\| \to 0$. Then by (3.8),

$$\begin{aligned} \big\|\big(\partial^\alpha x_j - \partial^\alpha f_j\big)\chi_{\Omega_j^{\alpha,0}}\big\| &\leqslant \left\| \left| \left(\frac{\partial^\alpha x_j + \partial^\alpha x_0}{2} - \partial^\alpha f_j \right) \chi_{\Omega_j^{\alpha,0}} \right| - \big| \big(\partial^\alpha x_j - \partial^\alpha f_j\big)\chi_{\Omega_j^{\alpha,0}} \big| \right\| \\ &\quad + \left\| \left(\frac{\partial^\alpha x_j + \partial^\alpha x_0}{2} - \partial^\alpha f_j \right) \chi_{\Omega_j^{\alpha,0}} \right\| \to 0. \end{aligned} \tag{3.9}$$

Now, by (3.7)–(3.9), we get $\|(\partial^\alpha x_0 - \partial^\alpha f_j)\chi_{\Omega_j^{\alpha,0}}\| \to 0$. Thus, in view of (3.8) and (3.9), for any α with $0 \leqslant |\alpha| \leqslant m$,

$$\begin{aligned} \big\|\partial^\alpha x_j - \partial^\alpha x_0\big\| &= \big\|\big(\partial^\alpha x_j - \partial^\alpha x_0\big)\chi_{\Omega_j^{\alpha,0}} + \big(\partial^\alpha x_j - \partial^\alpha x_0\big)\chi_{\Omega_n^{\alpha,1}}\big\| \\ &\leqslant \big\|\big(\partial^\alpha x_j - \partial^\alpha x_0\big)\chi_{\Omega_j^{\alpha,1}}\big\| + \big\|\big(\partial^\alpha x_j - \partial^\alpha f_j\big)\chi_{\Omega_j^{\alpha,0}}\big\| + \big\|\big(\partial^\alpha x_0 - \partial^\alpha f_j\big)\chi_{\Omega_j^{\alpha,0}}\big\| \to 0. \end{aligned}$$

Hence $\|x_j - x_0\|_{m,A} \to 0$. □

We finish by the following example of the convex subset K in $W_{m,A}$ for which we can apply the theorems of this section.

Example. Let fix a function $f \in W_{m,A}$ and define the set

$$K_m = \{p\colon\ p \text{ is a polynomial of degree} \leqslant m \text{ and } p \leqslant f\},$$

where $\leqslant$ is the order relation in Orlicz–Sobolev spaces, that is, for all derivatives $\partial^\alpha p(t) \leqslant \partial^\alpha f(t)$ a.e. in Ω, with $\alpha = (\alpha_1, \dots, \alpha_n)$ and $|\alpha| \leqslant m$. The set K_m is convex. According to Theorem 3.1 if A satisfies condition Δ_2 then if the element of best approximation to f in K exists then it is unique. Moreover, by Theorem 3.2, the minimal distance between f and K can be approximated by the minimizing Cauchy sequence $\{x_n\} \subset K$. So we can approximate given f by polynomials p of degree less than m in the sense that f and p as well as all their derivatives of degree $\leqslant m$ are close in the norm of Orlicz space L_A, and in particular if $A(u) = u^p$, $1 \leqslant p < \infty$, in the norm of L_p-space. In addition we have the advantage that p is pointwise (a.e. in Ω) less than f together with their appropriate derivatives.

Acknowledgments

We are grateful to our colleagues Grzegorz Lewicki and Leslaw Skrzypek for their valuable suggestions which improved the final version of the paper.

References

[1] R.A. Adams, Sobolev Spaces, Academic Press, New York, 1975.
[2] M.A. Akcoglu, L. Sucheston, On uniform monotonicity of norms and ergodic theorems in function spaces, Rend. Circ. Mat. 8 (2) (1985) 325–335.
[3] F. Deutsch, Best Approximation in Inner Product Spaces, Canad. Math. Soc., Springer, 2001.

[4] A. Benkirane, J.P. Gossez, Some approximation properties in Orlicz–Sobolev spaces, Studia Math. 74 (1) (1982) 17–24.
[5] G. Birkhoff, Lattice Theory, Amer. Math. Soc., Providence, RI, 1979.
[6] S.T. Chen, Geometry of Orlicz spaces, Dissertationes Math. (Rozprawy Mat.) 356 (1996).
[7] S.T. Chen, C.Y. Hu, On the extreme points and rotundity of Luxemburg norm on Orlicz–Sobolev spaces, Natur. Sci. J. Harbin Normal Univ. 17 (2) (2001) 1–6 (in Chinese).
[8] S.T. Chen, C.Y. Hu, X.J. Charles Zhao, Uniform rotundity of Orlicz–Sobolev spaces, Soochow J. Math. 29 (3) (2003) 299–312.
[9] Y.A. Cui, T.F. Wang, Uniform coefficients and application in Orlicz sequence spaces, Acta Math. Appl. Sin. 20 (3) (1997) 362–365 (in Chinese).
[10] X. Fan, J. Shen, D. Zhao, Sobolev embedding theorems for spaces $W_{k,p(x)}$, J. Math. Anal. Appl. 262 (2001) 749–760.
[11] J.P. Gossez, Orlicz–Sobolev spaces and nonlinear elliptic boundary value problems, in: Nonlinear Analysis, Function Spaces and Applications, Proc. Spring School, Horni Bradlo, 1978, pp. 59–94.
[12] J.P. Gossez, On a property of Orlicz–Sobolev spaces, in: Nonlinear Analysis and Optimization, Bologna, 1982, pp. 102–105.
[13] J.P. Gossez, On a property of Orlicz–Sobolev spaces, in: Trends in Theory and Practice of Differential Equations, Arlington, TX, 1982, in: Lect. Notes Pure Appl. Math., vol. 90, Dekker, New York, 1984, pp. 197–200.
[14] P. Harjulechto, P. Hasto, M. Koskenoja, S. Veronen, Sobolev capacity on the space $W_{1,p(\cdot)}(\mathbb{R}_n)$, J. Funct. Spaces Appl. 1 (2003) 17–33.
[15] H. Hudzik, Banach lattices with order isometric copies of l^∞, Indag. Math. (N.S.) 9 (4) (1998) 521–527.
[16] H. Hudzik, On density of $C_0^\infty(\Omega)$ in generalized Orlicz–Sobolev space $W_M^k(\Omega)$ for every open set $\Omega \subset \mathbb{R}^n$, Comment. Math. Prace Mat. 20 (1977) 65–78.
[17] H. Hudzik, On embedding theorems from Orlicz–Sobolev space $W_M^k(\Omega)$ into $C^m(\Omega)$ for every open set $\Omega \subset \mathbb{R}^n$, Comment. Math. Prace Mat. 20 (1978) 341–364.
[18] H. Hudzik, The problems of separability, duality, reflexivity and comparison for generalized Orlicz–Sobolev space $W_M^k(\Omega)$, Comment. Math. Prace Mat. 21 (1979) 315–324.
[19] H. Hudzik, A. Kamińska, Monotonicity properties of Lorentz spaces, Proc. Amer. Math. Soc. 123 (9) (1995) 2715–2721.
[20] H. Hudzik, A. Kamińska, M. Mastyło, Monotonicity and rotundity properties in Banach lattices, Rocky Mountain J. Math. 30 (3) (2000) 933–950.
[21] H. Hudzik, W. Kowalewski, G. Lewicki, Approximative compactness and full rotundity in Musielak–Orlicz spaces and Orlicz-Lorentz spaces, Z. Anal. Anwend. 25 (2006) 163–192.
[22] H. Hudzik, W. Kurc, Monotonicity properties of Musielak–Orlicz spaces and dominated best approximation in Banach lattices, J. Approx. Theory 95 (1998) 353–368.
[23] H. Hudzik, X.B. Liu, T.F. Wang, Points of monotonicity in Musielak–Orlicz function spaces endowed with the Luxemburg norm, Arch. Math. 82 (2004) 534–545.
[24] H. Hudzik, A. Narloch, Relationships between monotonicity and complex rotundity properties with some consequences, Math. Scand. 96 (2005) 289–306.
[25] S. Kilmer, W.M. Kozlowski, G. Lewicki, Best approximations in modular function spaces, J. Approx. Theory 63 (1990) 338–367.
[26] M.A. Krasnoselskii, Ya.B. Rutickii, Convex Functions and Orlicz Spaces, Noordhoff, Groningen, 1961.
[27] A. Kroó, J.J. Swetits, On uniqueness of best L_1-approximation in Sobolev spaces, Numer. Funct. Anal. Optim. 13 (1–2) (1992) 29–41.
[28] W. Kurc, Strictly and uniformly monotone Musielak–Orlicz spaces and applications to best approximation, J. Approx. Theory 69 (2) (1992) 173–187.
[29] D. Landers, R. Rogge, Best approximation in L_ϕ spaces, Z. Wahrscheinlichkeitstheorie Verv. Gebiete 51 (1980) 215–237.
[30] J. Lindenstrauss, L. Tzafriri, Classical Banach Spaces II, Springer-Verlag, Berlin/New York, 1979.
[31] X.B. Liu, T.F. Wang, Upper (lower) monotone coefficient of a point in Orlicz sequence space, J. Heilongjiang Univ. Natur. Sci. 18 (1) (2001) 1–4 (in Chinese).
[32] Y.M. Lü, J.M. Wang, T.F. Wang, Monotone coefficients and monotonicity of Orlicz spaces, Rev. Mat. Complut. 12 (1) (1999) 105–114.
[33] L. Maligranda, Orlicz Spaces and Interpolation, Seminars in Math., vol. 5, Univ. of Campinas, 1989.
[34] J. Musielak, Orlicz Spaces and Modular Spaces, Lecture Notes in Math., vol. 1034, Springer, Berlin, 1983.